Metadata For the Masses: Implications of the Pervasive, Easy Availability of Metadata in Text, Video, Photography and Objects

Shaun Foster

Rochester Institute of Technology
scffaa@rit.edu
www.growmultimedia.com

Jacob Brostoff

j.brostoff@gmail.com

Abstract
Metadata "is a set of data that describes and gives information about other data". The widespread availability of metadata is permeating nearly every facet of our lives. It is creating a new paradigm moving from device and file specific locations to access by "the cloud", using remote computer networks for storage and processing. Other metadata is being created from devices interpreting the real world and also from sensors embedded in various physical objects giving real time feedback. This paper focus on how metadata has impacted text, video, photography and objects allowing for faster and better communication, analysis and knowledge of the world.

Keywords
Metadata, ambient media, hypertext, text, photography, GPS

Metadata "is a set of data that describes and gives information about other data"[1]. Early non-digital metadata was used in library card catalogs, giving information such as name, location, and a summary of a books contents. The metadata would make it easy to decide if we wanted to look at this book or another. A similar application of metadata is used by digital audio, video and text files. Audio file metadata might contain the name of the artist, album, or length of a song. There has been a rapid growth of the availability and complexity of metadata provided from various information services.

Beyond simple growth in availability the nature of metadata is evolving in usage, location and type. Large amounts of metadata are used by computers without our direct knowledge in order to launch a correct program or structure web pages based on underlying metadata specifications. Location and structure of metadata have transformed from localized to a non-localized format. Previously all data files were placed on a hard drive inside the computer we were using. Metadata is changing into information that can be accessed from anywhere and longer exists in one specific location. New types of metadata are starting to permeate every aspect of our daily lives. We have sensor networks in our mobile devices, cars, GPS, etc. Being able to get nearly instant access to a large amount of meta-information has added a "sixth sense" of clarity and orientation. Our aim is to offer views into the growth and changes in metadata availability focusing within the fields text, video, physical objects and photography and make basic predictions for future metadata permutation and usage.

TEXT AND METADATA

There has been significant growth in text metadata. Currently text metadata includes font, point size, bold, italic, justification, location of file, number of words, etc. However, there is a growing availability of metadata within text used for improving the speed and emotional tone of communication.

Communication with text has changed. Previously when using written communication an assertion would be made, then defined and justified with long paragraphs of supporting information. Hyperlinks[2] are now used to streamline communication by instantly linking the viewer to support information. The reader can opt to follow the link, or not.. Another example from online collaborative paper writing (like this one) allows comments that are linked to areas being worked on helping speed critique, and clarification.

Visually text is also changing. Beyond embedding literal data into the text, there is more iconic representation of words using acronyms and emotions as emoticons. Emoticons and acronyms are changing and embedding what used to be mere text into semiotic metadata. Instead

of saying “I’m happy”, it is simpler to encode it into a symbol “:-)” or use “LOL” (for “laugh out loud”). There will undoubtedly be further augmentation of emotional or visual impact and meaning through additional rapid visual manipulation of text words. More efficiently changing font, size, bold, and italic for visual impact will encourage the growth of new visual systems in text. Beyond the visual, movement and audio are also being combined as metadata sources. New input technologies like the Microsoft pressure sensitive keyboard [3] could rapidly change the way we communicate with text by adding a layer of pressure metadata to text. Although this may initially be applied simply to text formatting or used in video games, there are other applications for adding movement to text as well. Currently there is a large amount of text animation done with motion graphics. But this could also be incorporated into text communication, adding a layer of kinesthetic metadata to text. The online service ToneCheck [5] analyzes text documents, mostly e-mails, to make sure the underlying emotional tone is e-mail friendly. Currently adding text by gesturing at a computer is considered humorous, but not too far away from becoming fact. Google’s 2011 April Fools video [4] played with this idea, yet within 48 hours it was a prototyped reality created by ICTxR Lab [6].

The emerging trends in metadata within text based communication show an evolution in communication. Metadata-embedded text will more clearly and quickly express ideas, feelings. It is easy to imagine a further synthesis of multi-sensory feedback within future text communications. Greater visual changes in text font size, shape and color, as well as movement, touch, and sound will change how we communicate.

ViDEO AND METADATA; AUTOMATIC CREATION FROM ANALOG SOURCES

Beyond the blooming growth of metadata there are also many programs being used to interpret and create metadata from analog sources.

The video editing software Adobe Premiere [7] creates text metadata inside of video. It does this by scanning the audio track and performing a voice-to-text analysis. With this information a user can search a video for specific keywords, going directly to the desired part of the video.

Autodesk MatchMover software [8] analyzes video footage generating a 3D visualization of the environment. This allows for the creation of a virtual 3D camera, allowing for compositing of 3D objects into moving video footage. This type of hybrid analog/digital metadata information from computer vision is being combined with directional and global positioning systems [9](GPS) data and being used in products like Google Goggles [10], augmented reality [11], and the new geolocation picture sharing iPhone app Color [12].

The future of video and automatic metadata creation will see further increases in automatic identification and labeling. Facial recognition, object recognition, costs comparisons. Video analysis technology will be further integrated with motion capture technology like the Xbox Kinect, to give critiques on athletic form, to more refined critiques like body posture and poise for business presentations.

LOCATION AND METADATA

Metadata is changing its relationship from local hard drive storage on a single computer to storage accessible from any networked computer. Decentralized storage of data is changing the way we work. In the past few years data has begun moving to virtual locations which are accessible from any networked location. Google Docs, blogs, email and web pages are the most common form of decentralized information. Impacts from rapid shifts in technology have been seen in the entertainment industry. Analog media such as audiotape and videotape once stored music and video. Then the storage changed to objects holding digital information, such as CDs and DVDs. Next video and music became more mobile and less about being stored in a directly-accessible physical medium, changing to audio-data files, MP3, WAV, MOV etc. One step further and now files are no longer stored on a local machine. Video and music can stream from a list to a user’s device. The bulk of information is stored non-locally. Netflix, Hulu and Amazon.com’s new “Cloud Player”[13] work by looking at users’ metadata list of their songs, artists, and music videos and then gets them streamed to their device on demand.

This is a major cultural change from a society which placed value in the physical to to a society that is more focused on being able to access data. This shift, “from atoms to bits”[14] is causing huge transformations in businesses formerly linked to making profit from the distribution of physical media such as CDs, newspapers, and DVDs. The newspaper, film and video game industries are all experiencing fundamental shifts in the way they do business. It is more convenient for most people to get a movie streamed to their TV than to go pick it up at a physical store.

DYNAMIC METADATA FROM PHYSICAL

OBJECTS

Text, music, and video files do not change when they are played multiple times or in different locations. Their data is static. Metadata gathered dynamically from physical objects is growing exponentially. Information that changes over time and is accessed on demand from and object physical objects will open the door to an enormous amount of new knowledge about ourselves and the way we live.

Cell phones are a common source of dynamic data. Triangulation and GPS information from cell phones can tell location information. Many cell phones can now also scan bar codes on objects in stores, such as food or books as well as augmented reality (AR) quick response (QR) codes. The bar code then links directly as in an AR code or or searches for the latest reviews, check calorie counts or even look to see if the item is on sale at another store. It is also possible to scan your airplane ticket to do an early check-in or see if your flight is on time. These are only part of a changeover from an analog world to a world laced with digital metadata. With the exception of the cell phone, the objects being scanned are for the most part non-dynamic, nor do these objects sense or emit metadata of their own. The greatest impact on culture from data and metadata will come from new types of objects that talk to us and each other, generating dynamic metadata.

Low cost sensors and remote frequency ID (RFID) tags are now embedded in many types of objects. These objects generate new, continuous and spontaneous streams of metadata . Pressure sensors for car tires detect inflation levels. Weight sensors in car seats can detect by the weight of the passenger who is driving and make seat adjustments. Sensors in bridges and other "smart" materials can confirm structural integrity. Our GPS-enabled cell phones can both tell us how to get somewhere, but also generate metadata about where were we and where we are going.

Future trends from objects that can instantly talk to other objects are many. Buying an object might be as simple as walking out of a store with it. Performing an inventory could be done in an instant. While initially the application and evaluation of new sources of dynamic metadata will undoubtedly be lined to commercial applications. GPS and RFID open up a culture for an hybridization of physical objects and being able to access them over the Internet. Object hyperlinking[15] and embedded real-time data opens up a whole new world of possibilities for an ontology of objects though various services like Pachube [16].

IMPLICATIONS OF THE PERVASIVE AVAILABILITY OF METADATA: AN EXAMPLE FROM PHOTOGRAPHY

Photography is both art and science. Skilled photographers know the limits of their equipment, and what kinds of camera settings will yield various kinds of results. Many of photography's parameters, like depth-of-field, are calculable. For others, such as the formula for avoiding camera shake, there are widely used rules-of-thumb. However, the ability to take a shot with engaging composition or compelling content remains part of the mystery-and mystique-of good photography.

Today's digital cameras capture information about nearly all of the measurable parameters in a photograph. This metadata includes information about focal length, aperture, shutter speed, film plane sensitivity, the lens and camera model used, distortion correction, image stabilization, white balance, metering pattern, and so on. When transferred to a computer, programs like Apple Aperture can sort photographs based on metadata about the entire digital corpus of a photographer. For example, Aperture [17] can find every photograph shot at a particular focal length, or every photograph shot with a flash.

While in the past, some photographers would consistently record this kind of information, others would not (or in the case of many street and sports photographers, could not) reliably capture every detail of how a shot was taken as part of the photographic process. Some information, such as film speed and camera model used, was often memorable or easily available. Other information, such as the precise focal length of a shot taken with a zoom lens, was nearly impossible to reliably capture in a rigorous way.

For contemporary photographers, the pervasive availability of many kinds of metadata has profound implications.

Being able to analyze a photographic corpus based on pervasively available metadata can add additional layers of quantification and understanding of the patterns in the corpus. It is now easy to quantify how many photographs were shot with a given aperture or shutter speed. This, combined with the ability to rate the quality of a given photograph (another of Aperture's features) enables the photographer to ask more sophisticated questions about her corpus and practices, such as: How many highly-rated photographs did I shoot with my 20mm lens wide open?

CONCLUSIONS

Old and new kinds of metadata are more widely available and easily accessible now than they have ever been. Having easy access to data about data has increased the culture's awareness of new uses for existing data. It has also created new possibilities for analyzing data that was previously tedious or impractical to use in a meaningful way. In photography, the pervasive and easy availability of many kinds of metadata means that photographers have new opportunities to catalog and analyze their body of work, and potentially to modify their practice. Metadata will no doubt change how other disciplines make use of their existing data in more and more sophisticated and self-aware ways.

REFERENCES

1. Metadata Definition: http://www.wto.org/english/theWTO_e/glossary_e/metadata_e.htm
2. Hyperlink Definition: http://en.wikipedia.org/wiki/Hyperlink
3. Microsoft Pressure Sensitive Keyboard: http://gizmodo.com/#!5331577/microsofts-pressure+sensitive-keyboard-could-change-the-way-you-type
4. Google Motion: http://mail.google.com/mail/help/motion.html
5. Emotional Tone Check: http://www.tonecheck.com/
6. ICTxR Labs Motion: http://www.youtube.com/watch?v=Lfso7_i9Ko8&feature=player_embedded
7. Adobe Premiere: http://help.adobe.com/en_US/PremierePro/4.0/WS29F692D2-7A49-44db-A109-AE016B766767.html
8. AutodeskMatchmover: http://download.autodesk.com/us/maya/2011help-matchmover/index.html
9. GPS Definition: http://en.wikipedia.org/wiki/Global_Positioning_System
10. Google Goggles: http://www.google.com/mobile/goggles/#text
11. Augmented Reality Definition: http://en.wikipedia.org/wiki/Augmented_reality
12. Color App: http://blogs.forbes.com/andygreenberg/2011/03/28/color-app-hack-lets-you-spy-on-anyones-photos-anywhere/
13. Amazon Cloud Player: http://www.engadget.com/2011/03/29/amazon-cloud-player-goes-live-streams-music-on-your-computer-an/
14. Negroponte, Nicholas *Being Digital*, New York Random House 1995, P11
15. Object Hyperlinking Definition: http://en.wikipedia.org/wiki/Object_hyperlinking
16. "Network of things" Pachube: http://www.pachube.com/
17. Apple Aperture: http://www.apple.com/aperture/